

PCT09

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/926,299

DATE: 10/23/2001
TIME: 15:59:48

Input Set : A:\212289US0PCT.txt
Output Set: N:\CRF3\10232001\I926299.raw

PS
ENTERED

3 <110> APPLICANT: GUNJI, YOSHIYA
4 YASUEDA, HISASHI
5 SUGIMOTO, SHINICHI
6 TSUJIMOTO, NOBUHARU
7 SHIMAOKA, MEGUMI
8 MIYATA, YURI
9 OBA, MANAMI
11 <120> TITLE OF INVENTION: L-AMINO ACID-PRODUCING BACTERIUM AND METHOD FOR PRODUCING L-AMINO ACID
13 <130> FILE REFERENCE: 212289US0PCT
C--> 15 <140> CURRENT APPLICATION NUMBER: US/09/926,299
C--> 15 <141> CURRENT FILING DATE: 2001-10-09
15 <150> PRIOR APPLICATION NUMBER: PCT/JP 00/02295
16 <151> PRIOR FILING DATE: 2000-04-17
18 <150> PRIOR APPLICATION NUMBER: JP 11-103143
19 <151> PRIOR FILING DATE: 1999-04-09
21 <150> PRIOR APPLICATION NUMBER: JP 11-169447
22 <151> PRIOR FILING DATE: 1999-06-16
24 <150> PRIOR APPLICATION NUMBER: JP 11-368097
25 <151> PRIOR FILING DATE: 1999-12-24
27 <160> NUMBER OF SEQ ID NOS: 20
29 <170> SOFTWARE: PatentIn version 3.1
31 <210> SEQ ID NO: 1
32 <211> LENGTH: 1197
33 <212> TYPE: DNA
34 <213> ORGANISM: Escherichia coli
36 <220> FEATURE:
37 <221> NAME/KEY: CDS
38 <222> LOCATION: (272)..(1147)
39 <223> OTHER INFORMATION:
42 <400> SEQUENCE: 1
43 ccaggcgact gtcttcaata ttacagccgc aactactgac atgacgggtg atggtgttca 60
45 caattccacg gcgatcggca cccaacgcag tgatcaccag ataatgtgtt gcgatgacag 120
47 tgtcaaaactg gttattcctt taaggggtga gttgttctta aggaaagcat aaaaaaaca 180
49 tgcatacaac aatcagaacg gttctgtctg cttgctttta atgccatacc aaacgtacca 240
51 ttgagacact tgtttgaca gaggatggcc c atg ttc acg gga agt att gtc 292
52 Met Phe Thr Gly Ser Ile Val
53 1 5
55 gcg att gtt act ccg atg gat gaa aaa ggt aat gtc tgt cgg gct agc 340
56 Ala Ile Val Thr Pro Met Asp Glu Lys Gly Asn Val Cys Arg Ala Ser
57 10 15 20
59 ttg aaa aaa ctg att gat tat cat gtc gcc agc ggt act tcg gcg atc 388
60 Leu Lys Lys Leu Ile Asp Tyr His Val Ala Ser Gly Thr Ser Ala Ile
61 25 30 35
63 gtt tct gtt ggc acc act ggc gag tcc gct acc tta aat cat gac gaa 436
64 Val Ser Val Gly Thr Thr Gly Glu Ser Ala Thr Leu Asn His Asp Glu
65 40 45 50 55
67 cat gct gat gtg gtg atg atg acg ctg gat ctg gct gat ggg cgc att 484

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/926,299

DATE: 10/23/2001

TIME: 15:59:48

Input Set : A:\212289USOPCT.txt

Output Set: N:\CRF3\10232001\I926299.raw

```

68 His Ala Asp Val Val Met Met Thr Leu Asp Leu Ala Asp Gly Arg Ile
69          60          65          70
71 ccg gta att gcc ggg acc ggc gct aac gct act gcg gaa gcc att agc      532
72 Pro Val Ile Ala Gly Thr Gly Ala Asn Ala Thr Ala Glu Ala Ile Ser
73          75          80          85
75 ctg acg cag cgc ttc aat gac agt ggt atc gtc ggc tgc ctg acg gta      580
76 Leu Thr Gln Arg Phe Asn Asp Ser Gly Ile Val Gly Cys Leu Thr Val
77          90          95          100
79 acc cct tac tac aat cgt ccg tgc caa gaa ggt ttg tat cag cat ttc      628
80 Thr Pro Tyr Tyr Asn Arg Pro Ser Gln Glu Gly Leu Tyr Gln His Phe
81          105          110          115
83 aaa gcc atc gct gag cat act gac ctg ccg caa att ctg tat aat gtg      676
84 Lys Ala Ile Ala Glu His Thr Asp Leu Pro Gln Ile Leu Tyr Asn Val
85 120          125          130          135
87 ccg tcc cgt act ggc tgc gat ctg ctc ccg gaa acg gtg ggc cgt ctg      724
88 Pro Ser Arg Thr Gly Cys Asp Leu Leu Pro Glu Thr Val Gly Arg Leu
89          140          145          150
91 gcg aaa gta aaa aat att atc gga atc aaa gag gca aca ggg aac tta      772
92 Ala Lys Val Lys Asn Ile Ile Gly Ile Lys Glu Ala Thr Gly Asn Leu
93          155          160          165
95 acg cgt gta aac cag atc aaa gag ctg gtt tca gat gat ttt gtt ctg      820
96 Thr Arg Val Asn Gln Ile Lys Glu Leu Val Ser Asp Asp Phe Val Leu
97          170          175          180
99 ctg agc ggc gat gat gcg agc gcg ctg gac ttc atg caa ttg ggc ggt      868
100 Leu Ser Gly Asp Asp Ala Ser Ala Leu Asp Phe Met Gln Leu Gly Gly
101          185          190          195
103 cat ggg gtt att tcc gtt acg act aac gtc gca gcg cgt gat atg gcc      916
104 His Gly Val Ile Ser Val Thr Thr Asn Val Ala Ala Arg Asp Met Ala
105 200          205          210          215
107 cag atg tgc aaa ctg gca gca gaa gaa cat ttt gcc gag gca cgc gtt      964
108 Gln Met Cys Lys Leu Ala Ala Glu Glu His Phe Ala Glu Ala Arg Val
109          220          225          230
111 att aat cag cgt ctg atg cca tta cac aac aaa cta ttt gtc gaa ccc      1012
112 Ile Asn Gln Arg Leu Met Pro Leu His Asn Lys Leu Phe Val Glu Pro
113          235          240          245
115 aat cca atc ccg gtg aaa tgg gca tgt aag gaa ctg ggt ctt gtg gcg      1060
116 Asn Pro Ile Pro Val Lys Trp Ala Cys Lys Glu Leu Gly Leu Val Ala
117          250          255          260
119 acc gat acg ctg cgc ctg cca atg aca cca atc acc gac agt ggt cgt      1108
120 Thr Asp Thr Leu Arg Leu Pro Met Thr Pro Ile Thr Asp Ser Gly Arg
121          265          270          275
123 gag acg gtc aga gcg gcg ctt aag cat gcc ggt ttg ctg taaagtttag      1157
124 Glu Thr Val Arg Ala Ala Leu Lys His Ala Gly Leu Leu
125 280          285          290
127 ggagatttga tggcttactc tgttcaaaag tcgcgcctgg      1197
130 <210> SEQ ID NO: 2
131 <211> LENGTH: 292
132 <212> TYPE: PRT
133 <213> ORGANISM: Escherichia coli

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/926,299

DATE: 10/23/2001
 TIME: 15:59:48

Input Set : A:\212289US0PCT.txt
 Output Set: N:\CRF3\10232001\I926299.raw

135 <400> SEQUENCE: 2
 137 Met Phe Thr Gly Ser Ile Val Ala Ile Val Thr Pro Met Asp Glu Lys
 138 1 5 10 15
 141 Gly Asn Val Cys Arg Ala Ser Leu Lys Lys Leu Ile Asp Tyr His Val
 142 20 25 30
 145 Ala Ser Gly Thr Ser Ala Ile Val Ser Val Gly Thr Thr Gly Glu Ser
 146 35 40 45
 149 Ala Thr Leu Asn His Asp Glu His Ala Asp Val Val Met Met Thr Leu
 150 50 55 60
 153 Asp Leu Ala Asp Gly Arg Ile Pro Val Ile Ala Gly Thr Gly Ala Asn
 154 65 70 75 80
 157 Ala Thr Ala Glu Ala Ile Ser Leu Thr Gln Arg Phe Asn Asp Ser Gly
 158 85 90 95
 161 Ile Val Gly Cys Leu Thr Val Thr Pro Tyr Tyr Asn Arg Pro Ser Gln
 162 100 105 110
 165 Glu Gly Leu Tyr Gln His Phe Lys Ala Ile Ala Glu His Thr Asp Leu
 166 115 120 125
 169 Pro Gln Ile Leu Tyr Asn Val Pro Ser Arg Thr Gly Cys Asp Leu Leu
 170 130 135 140
 173 Pro Glu Thr Val Gly Arg Leu Ala Lys Val Lys Asn Ile Ile Gly Ile
 174 145 150 155 160
 177 Lys Glu Ala Thr Gly Asn Leu Thr Arg Val Asn Gln Ile Lys Glu Leu
 178 165 170 175
 181 Val Ser Asp Asp Phe Val Leu Leu Ser Gly Asp Asp Ala Ser Ala Leu
 182 180 185 190
 185 Asp Phe Met Gln Leu Gly Gly His Gly Val Ile Ser Val Thr Thr Asn
 186 195 200 205
 189 Val Ala Ala Arg Asp Met Ala Gln Met Cys Lys Leu Ala Ala Glu Glu
 190 210 215 220
 193 His Phe Ala Glu Ala Arg Val Ile Asn Gln Arg Leu Met Pro Leu His
 194 225 230 235 240
 197 Asn Lys Leu Phe Val Glu Pro Asn Pro Ile Pro Val Lys Trp Ala Cys
 198 245 250 255
 201 Lys Glu Leu Gly Leu Val Ala Thr Asp Thr Leu Arg Leu Pro Met Thr
 202 260 265 270
 205 Pro Ile Thr Asp Ser Gly Arg Glu Thr Val Arg Ala Ala Leu Lys His
 206 275 280 285
 209 Ala Gly Leu Leu
 210 290
 213 <210> SEQ ID NO: 3
 214 <211> LENGTH: 2147
 215 <212> TYPE: DNA
 216 <213> ORGANISM: Escherichia coli
 218 <220> FEATURE:
 219 <221> NAME/KEY: CDS
 220 <222> LOCATION: (584)..(1930)
 221 <223> OTHER INFORMATION:
 224 <400> SEQUENCE: 3
 225 tcgaagtgtt tctgtagtgc ctgccaggca gcggtctgcg ttggattgat gtttttcatt 60

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/926,299

DATE: 10/23/2001
TIME: 15:59:48

Input Set : A:\212289US0PCT.txt
Output Set: N:\CRF3\10232001\I926299.raw

```

227 agcaatactc ttctgatttt gagaattgtg actttggaag attgtagcgc cagtcacaga 120
229 aaaatgtgat ggtttttagtg ccgttagcgt aatgttgagt gtaaaccott agcgagtgga 180
231 agcattttatt agctgaacta ctgaccgcca ggagtggatg aaaaatccgc atgaccccat 240
233 cgttgacaac cgccccgctc accctttatt tataaatgta ctacctgcgc tagcgagggc 300
235 cagaagaggc gcgttgccca agtaacgggt ttggaggagc cagtcctgtg ataacacctg 360
237 aggggggtgca tcgccgaggt gattgaacgg ctggccacgt tcatcatcgg ctaagggggc 420
239 tgaatccctt gggttgtcac cagaagcgtt cgcagtcggg cgtttcgcaa gtgggtggagc 480
241 actttctgggt gaaaatagta gcgaagtatc gctctgcgcc caccctcttt ccgctcttcc 540
243 cttgtgccaa ggctgaaaat ggatcccttg acacgaggta gtt atg tct gaa att 595
244                                     Met Ser Glu Ile
245                                     1
247 gtt gtc tcc aaa ttt ggc ggt acc agc gta gct gat ttt gac gcc atg 643
248 Val Val Ser Lys Phe Gly Gly Thr Ser Val Ala Asp Phe Asp Ala Met
249 5                                     10 15 20
251 aac cgc agc gct gat att gtg ctt tct gat gcc aac gtg cgt tta gtt 691
252 Asn Arg Ser Ala Asp Ile Val Leu Ser Asp Ala Asn Val Arg Leu Val
253 25 30 35
255 gtc ctc tcg gct tct gct ggt atc act aat ctg ctg gtc gct tta gct 739
256 Val Leu Ser Ala Ser Ala Gly Ile Thr Asn Leu Leu Val Ala Leu Ala
257 40 45 50
259 gaa gga ctg gaa cct ggc gag cga ttc gaa aaa ctc gac gct atc cgc 787
260 Glu Gly Leu Glu Pro Gly Glu Arg Phe Glu Lys Leu Asp Ala Ile Arg
261 55 60 65
263 aac atc cag ttt gcc att ctg gaa cgt ctg cgt tac ccg aac gtt atc 835
264 Asn Ile Gln Phe Ala Ile Leu Glu Arg Leu Arg Tyr Pro Asn Val Ile
265 70 75 80
267 cgt gaa gag att gaa cgt ctg ctg gag aac att act gtt ctg gca gaa 883
268 Arg Glu Glu Ile Glu Arg Leu Leu Glu Asn Ile Thr Val Leu Ala Glu
269 85 90 95
271 gcg gcg gcg ctg gca acg tct ccg gcg ctg aca gat gag ctg gtc agc 931
272 Ala Ala Ala Leu Ala Thr Ser Pro Ala Leu Thr Asp Glu Leu Val Ser
273 105 110 115
275 cac ggc gag ctg atg tcg acc ctg ctg ttt gtt gag atc ctg cgc gaa 979
276 His Gly Glu Leu Met Ser Thr Leu Leu Phe Val Glu Ile Leu Arg Glu
277 120 125 130
279 cgc gat gtt cag gca cag tgg ttt gat gta cgt aaa gtg atg cgt acc 1027
280 Arg Asp Val Gln Ala Gln Trp Phe Asp Val Arg Lys Val Met Arg Thr
281 135 140 145
283 aac gac cga ttt ggt cgt gca gag cca gat ata gcc gcg ctg gcg gaa 1075
284 Asn Asp Arg Phe Gly Arg Ala Glu Pro Asp Ile Ala Ala Leu Ala Glu
285 150 155 160
287 ctg gcc gcg ctg cag ctg ctc cca cgt ctc aat gaa ggc tta gtg atc 1123
288 Leu Ala Ala Leu Gln Leu Leu Pro Arg Leu Asn Glu Gly Leu Val Ile
289 165 170 175
291 acc cag gga ttt atc ggt agc gaa aat aaa ggt cgt aca acg acg ctt 1171
292 Thr Gln Gly Phe Ile Gly Ser Glu Asn Lys Gly Arg Thr Thr Thr Leu
293 185 190 195
295 ggc cgt gga ggc agc gat tat acg gca gcc ttg ctg gcg gag gct tta 1219
296 Gly Arg Gly Gly Ser Asp Tyr Thr Ala Ala Leu Leu Ala Glu Ala Leu

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/926,299

DATE: 10/23/2001
TIME: 15:59:48

Input Set : A:\212289US0PCT.txt
Output Set : N:\CRF3\10232001\I926299.raw

```

297          200          205          210          1267
299 cac gca tct cgt gtt gat atc tgg acc gac gtc ccg ggc atc tac acc
300 His Ala Ser Arg Val Asp Ile Trp Thr Asp Val Pro Gly Ile Tyr Thr
301          215          220          225
303 acc gat cca cgc gta gtt tcc gca gca aaa cgc att gat gaa atc gcg
304 Thr Asp Pro Arg Val Val Ser Ala Ala Lys Arg Ile Asp Glu Ile Ala
305          230          235          240          1315
307 ttt gcc gaa gcg gca gag atg gca act ttt ggt gca aaa gta ctg cat
308 Phe Ala Glu Ala Ala Glu Met Ala Thr Phe Gly Ala Lys Val Leu His
309 245          250          255          260          1363
311 ccg gca acg ttg cta ccc gca gta cgc agc gat atc ccg gtc ttt gtc
312 Pro Ala Thr Leu Leu Pro Ala Val Arg Ser Asp Ile Pro Val Phe Val
313          265          270          275          1411
315 ggc tcc agc aaa gac cca cgc gca ggt ggt acg ctg gtg tgc aat aaa
316 Gly Ser Ser Lys Asp Pro Arg Ala Gly Gly Thr Leu Val Cys Asn Lys
317          280          285          290          1459
319 act gaa aat ccg ccg ctg ttc cgc gct ctg gcg ctt cgt cgc aat cag
320 Thr Glu Asn Pro Pro Leu Phe Arg Ala Leu Ala Leu Arg Arg Asn Gln
321          295          300          305          1507
323 act ctg ctc act ttg cac agc ctg aat atg ctg cat tct cgc ggt ttc
324 Thr Leu Leu Thr Leu His Ser Leu Asn Met Leu His Ser Arg Gly Phe
325          310          315          320          1555
327 ctc gcg gaa gtt ttc ggc atc ctc gcg ccg cat aat att tcg gta gac
328 Leu Ala Glu Val Phe Gly Ile Leu Ala Arg His Asn Ile Ser Val Asp
329 325          330          335          340          1603
331 tta atc acc acg tca gaa gtg agc gtg gca tta acc ctt gat acc acc
332 Leu Ile Thr Thr Ser Glu Val Ser Val Ala Leu Thr Leu Asp Thr Thr
333          345          350          355          1651
335 ggt tca acc tcc act ggc gat acg ttg ctg acg caa tct ctg ctg atg
336 Gly Ser Thr Ser Thr Gly Asp Thr Leu Leu Thr Gln Ser Leu Leu Met
337          360          365          370          1699
339 gag ctt tcc gca ctg tgt cgc gtg gag gtg gaa gaa ggt ctg gcg ctg
340 Glu Leu Ser Ala Leu Cys Arg Val Glu Val Glu Glu Gly Leu Ala Leu
341          375          380          385          1747
343 gtc gcg ttg att ggc aat gac ctg tca aaa gcc tgc ggc gtt ggc aaa
344 Val Ala Leu Ile Gly Asn Asp Leu Ser Lys Ala Cys Gly Val Gly Lys
345          390          395          400          1795
347 gag gta ttc ggc gta ctg gaa ccg ttc aac att cgc atg att tgt tat
348 Glu Val Phe Gly Val Leu Glu Pro Phe Asn Ile Arg Met Ile Cys Tyr
349 405          410          415          420          1843
351 ggc gca tcc agc cat aac ctg tgc ttc ctg gtg ccc ggc gaa gat gcc
352 Gly Ala Ser Ser His Asn Leu Cys Phe Leu Val Pro Gly Glu Asp Ala
353          425          430          435          1891
355 gag cag gtg gtg caa aaa ctg cat agt aat ttg ttt gag taaatactgt
356 Glu Gln Val Val Gln Lys Leu His Ser Asn Leu Phe Glu
357          440          445          1940
359 atggcctgga agctatatatt cgggccgtat tgattttctt gtcaactatgc tcatcaataa
361 acgagcctgt actctgttaa ccagcgtctt tatcgagaa taattgcctt taattttttt
363 atctgcatct ctaattaatt atcgaaagag ataaatagtt aagagaaggc aaaatgaata

```

Use of n and / or Xaa has been detected in the
Sequence Listing. Review the Sequence Listing
to ensure a corresponding explanation is present
in the <220> to <223> fields of each sequence
using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/926,299

DATE: 10/23/2001

TIME: 15:59:49

Input Set : A:\212289USOPCT.txt

Output Set: N:\CRF3\10232001\I926299.raw

L:15 M:270 C: Current Application Number differs, Replaced Current Application No
L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:831 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:832 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:949 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1639 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13